

Rational Choice Theory: A Forum for Exchange of Ideas between the Hard and Social Sciences in Predictive Behavioral Modeling

Sun-Ki Chai[†]

[†] sunki@hawaii.edu, Department of Sociology, University of Hawai`i, Honolulu HI

Abstract The rational choice model of human behavior provides general assumptions that underlie most of the predictive behavioral modeling done in the social sciences. Surprisingly, despite its origins in the work of eminent mathematicians and computer scientists and its current prominence in the social sciences, there has been relatively little interest among hard scientists in incorporating rational choice assumptions into their agent-based analysis of behavior. It is argued here that doing so will introduce greater theoretical generality into agent-based models, and will also provide hard scientists an opportunity to contribute solutions to known weaknesses in the conventional version of rational choice. This in turn will invigorate the dialogue between social scientists and hard scientists studying social phenomena.

1 Introduction

It is very difficult to survey recent work within the social sciences on social computing, behavioral modeling, and prediction in any reasonably concise fashion. For many decades, mathematical, predictive models of behavior have characterized almost all of economic theory and major parts of other social science disciplines as well, particularly within the broad framework of the theoretical approach known as rational choice theory. So in a sense, formal work on predictive behavioral modeling makes up a large portion of the social sciences as a whole.

Rather than attempting to cover such an expanse of research, I will limit myself to investigating the base set of assumptions about human nature that underlies much if not most of this social science literature, the assumptions of rational choice. The next section of this paper will discuss the origins of the rational choice model during an era in which the social sciences were heavily influenced by developments in the “hard” sciences. Following this, I will discuss debates on the strengths and weaknesses of the rational choice model, and attempts to improve it. In the concluding section, I will argue that this debate over rational choice presents an opportunity for hard scientists studying human behavior to improve the generality of their own models while once again contributing to the development of the social sciences.

2 The Origins of the Rational Choice Model in Mathematics and Computer Science

The use of hard science methods as a basis for the prediction of behavior in the social sciences is nothing new. For the leaders of the “marginalist” revolution of the late 19th century that led the modern form of mathematical economics, physics provided a paradigm to be emulated (Mirosky, 1991). While late 19th century sociologists were more circumspect about adopting scientific over humanistic methods, as typified by Weber’s concept of *verstehen*, the general desire to build “science of society” was widely accepted (Lepenies and Hollingdale, 1988).

More concretely, and of greater direct relevance to the issues of this conference, foundational work in expected utility theory and game theory came about as a collaboration between the computer scientist/mathematician John von Neumann and the economist Oskar Morgenstern (1944). Their basic formal frameworks, augmented by the later work of Leonard Savage, John Nash (both mathematicians), and others, have provided the technical apparatus, as well as some of the basic assumptions, for most mathematical predictive theories of behavior in the social sciences, which in turn go under the general rubric of formal rational choice theory.

However, despite the clear heritage of formal rational choice modeling in mathematics and computer science, subsequent development of the rational choice literature has not continued to join the social and hard sciences, but more often to separate the two. The bulk of subsequent development in the rational choice approach has taken place in the field of economics, with contributions from political science, sociology, and occasionally other related social sciences. Dialogue and collaboration with the hard sciences has been much less common than one might expect given the origins of the theoretical approach.

One major exception is evolutionary biology, where the pioneering work of John Maynard Smith (1982) triggered a fruitful dialog with economists that helped turned evolutionary game theory (Weibull, 1997) into a truly interdisciplinary field of technical research, and generated within economics such literatures as the evolutionary theory of the firm (Nelson and Winters, 2006 [1988]) and the theory of learning in organizations (Fudenberg and Levine, 1998).

Another exception is operations research, where game theory has long stood beside linear and linear programming, queueing theory, etc. as one of the basic palette of analytical techniques in the field (Parthasarathy et al, 1997). However, much of the interest in game theory within operations research has been in cooperative rather than non-cooperative game theory, which has somewhat reduced its impact on the social sciences. Moreover, it has been argued that while game theory has penetrated the conversation of OR applied scholars, less emphasis has been placed on producing foundational theoretical research (Shubik, 2002).

A specific example of early cross-fertilization is Robert Axelrod’s pioneering use of computer simulations to investigate the relative fitness of decision strategies in iterated Prisoner’s Dilemma games (Axelrod, 2005 [1984]; 1997). More

recently, there has been a burgeoning in the importation of ideas from the hard sciences, particularly computer science, to enrich social science mathematical methods. The use of agent-based social simulations and other techniques arising from artificial intelligence/cognitive science has become widely accepted in mainstream sociology (Bainbridge et al, 1994; Macy and Willer, 2002) and is making inroads into economics, particularly under rubric of behavioral economics (Camerer et al, 2003), which not incidentally takes Herbert Simon as one of its founding fathers.

3 Debate over Rational Choice within the Social Sciences

To some extent, this renewal of interest in models from the hard sciences is a result of frustration with the weaknesses of the conventional version of the rational choice model, whose assumptions about human behavior have dominated formal behavioral modeling in the social sciences. Its base assumptions can be summarized informally as the idea that an actor seeks to maximize preferences, usually reflected by utility function, a quantity that reflects the desirableness of conceivable outcomes, and does so in light of certain beliefs that she/he holds, which are usually represented as subjective probabilities, sometimes contingent, about the state of the environment. The existence of a utility function reflects certain inherent assumptions about the properties of preferences, most notably that they constitute a strict order having the properties of completeness/connectedness, irreflexivity, acyclicity, and transitivity. Beliefs are constrained to be consistent with each other and the usual axioms of probability theory (Chai, 1997; 2001, chap. 1).

In practice, these base assumptions are not enough to generate predictions about behavior, and must be augmented with auxiliary assumptions or methodologies that determine the substance of preferences and beliefs. In the conventional version of rational choice, preferences are assumed to be unidimensional, self-regarding, materialistic, and isomorphic across actors. In other words, preferences for all individuals can be represented by their own endowment of a single material good, typically money. Other typical additions to the base assumptions involve the utility attached to uncertain outcomes (expected utility) and the way that utility is cumulated over time (exponential discounting of future utility).

Also in the conventional version, beliefs are based on what can be called an “information assumption”, i.e. that beliefs derive solely from facts that an actor can observe directly, as well as propositions that can be inferred from such facts through logic or probability theory. The information assumption is a formal device that allows one to infer the totality of an actors beliefs from their past and present position (defined both spatially and socially) in the environment. One other addition to belief assumptions is common knowledge of rationality, the idea that actors are all aware of each other’s rationality (in the sense being described

here), a crucial assumption for generating predictions from the game theory formalism.

The great strength of conventional rational choice can be summarized by its combination of generality, decisiveness, parsimony, and relative accuracy. Generality refers to the applicability of the conventional rational choice model to virtually any environment and choice situation. Because its assumptions encompass the necessary characteristics for representing choice using game theory and expected utility formalizations, it can be applied to choice situation that can be represented using these formalizations.

In most such environments, conventional rational choice assumptions not only can be applied, but will generate predictions that specify a set of possible outcomes that is significantly more constrained than the set of conceivable outcomes. This is the quality of decisiveness, and it ensures that across these environments, predictions will not only tend to be falsifiable (as long as the outcomes represented are operationalizable), but also non-trivial, ruling out outcomes that are not *ipso facto* impossible or improbable.

The conventional assumptions, while more complex than the base ones, are parsimonious in that only a small number of variables that have to be considered in generating predictions about behavior. Indeed, actor preferences and beliefs are both endogenous to (specified from within) the conventional model, hence the only information that is needed to generate behavioral predictions within any environment are the past and present positions of the actor within the environment.

Finally, the predictions that conventional rational choice model has generated, despite the seeming handicap of carrying a single set of assumptions across the full domain of choice environments, has often generated predictions that are often both provocative and confirmed by empirical evidence. While most of these predictions have been its “home” environment of economic behavior, recent efforts have greatly expanded the realm of prediction to areas far afield. The recent phenomenon of “freakonomics” (Leavitt and Dubner, 2006) is simply the most recent manifestation of a long process of expansion in the domain of rational choice, a process that has been referred to most commonly, by both proponents and opponents, as “economic imperialism” (Radnitzky and Bernholz, 1987; Lazear, 2000).

No other existing major theoretical approach equals conventional rational choice in meeting this combination of criteria, and this has been seen as justifying the approach’s preeminence in social science predictive behavioral modeling. Put together, the criteria ensure that findings arising from the application of the conventional rational choice model cumulate into law-like statements based on an internally consistent view of human nature and behavior. Because of this cumulatibility, it is relatively easy to generalize from rational choice findings to formulate new, testable hypotheses in arenas where the model has yet to be applied, an important factor when the model is the basis for predictive technology that will may be used under future, unforeseen circumstances. Furthermore, because these hypotheses arise from the application of a general model that has been confirmed under a multiplicity of environments, there is greater reason to believe that

the hypotheses will be accurate than if they were simply formulated in an *ad hoc* fashion. As such, it is without contest the preeminent embodiment in the social sciences of the nomothetic method associated with the natural sciences (Windelband, 1894).

Despite these positive qualities, the conventional rational choice approach has over the past decade been subject to a host of criticisms that have exposed its major weaknesses in the harshest light. While it has always been accepted that the conventional assumptions comprise an extreme simplification of human psychology, the true extent to which this hampers its ability to generate accurate predictions has become increasingly clear.

The approach been known to generate major predictive anomalies across a wide set of environments. These include anomalies of indecisiveness, where theories based on conventional rational choice cannot generate predictions that substantially reduce the set of all feasible behavioral outcomes because the actors involved do not have sufficient information to decide from logical alone which choices are better than others. This has been found most frequently in environments where actors are engaged in repeated interactions of uncertain duration (Fudenberg and Maskin, 1986) or where there is a need for actors to coordinate behaviors with one another. They also include anomalies of inaccuracy, where conventional theories predict outcomes that differ systematically from observed behavior. This is found most often in “social dilemma” conditions where cooperation is mutually beneficial but cannot be enforced by individual reward (Henrich et al, 2004; Gintis et al, 2006). These anomalies are seen as being linked to underlying shortcomings with the motivational and cognitive assumptions of the conventional approach.

There have been two basic policies put forward to addressing these shortcomings. The first has been to advocate abandoning the rational choice approach altogether. The second has been to try keeping the baby while throwing out the bathwater, attempting to retain the strengths of the conventional approach while improving its predictive ability by augmenting it with new assumptions and techniques. In particular, the focus for the latter has been on replacing or augmenting its assumptions. This in turn can be done by allowing for greater variations in the types of beliefs and values that actors can hold, as well as allowing for change over time in these attributes of actors, often reflecting larger changes in culture. However, this must be done in a general, rather than *ad hoc*, manner in order to retain the strengths of the model.

In effect, any attempt to do so will constitute an attempt to insert culture into rational choice. This is not a simple task, and in the past it was taken for granted that the cultural and rational choice analysis were inherently incompatible (Harsanyi, 1969; Barry, 1970). While this is no longer accepted to be the case, and the way in culture can be inserted into rational choice are quite varied, it is much more difficult to do so in a way that retains the generality and decisiveness of the conventional approach.

There are arguably two primary ways of incorporating culture in such a general and decisive fashion. The first is the creation of general cultural dimensions and typologies that limit the kinds of cultural configurations that exist within a society. The second is an general endogenous theory of culture formation that can predict variations in preferences and beliefs across individuals and over time.

An initial attempt to combine both types of methods is the grid-group/coherence approach. One component makes use of the grid-group cultural dimensions developed by the eminent social anthropologist Mary Douglas (1970; 1984), adapting them to make them compatible with the base assumptions of rational choice analysis of behavior (Chai and Wildavsky 1994; Chai and Swedlow 1998). The other component introduces a coherence model of preference and belief formation that is based on the synthesis and formalization of various prominent theories of self found in psychology, sociology, and anthropology (Chai, 2001). However, this approach certainly is not the only possible way of incorporating cultural in a general fashion, and other efforts are awaited.

4 An Opportunity for Synergy in Reformulating Rational Choice

As the papers presented at this conference attest, there is no doubt that behavioral modeling has enjoyed a productive boom among scholars in the hard sciences, particularly in the fields of computer science, electrical and industrial engineering, and biology. This work has brought a new level of technical rigor to the formal analysis of behavior, as well as development of an array of technical tools for analyzing and predicting behavior.

Perhaps not surprisingly, the dominant mode of analysis for computer scientists, engineers, mathematicians, and biologists engaging in prediction of human behavior has been computational agent-based modeling. Somewhat less expectedly, hard scientists have tended put little effort into coming to some agreement about the basic general assumptions regarding human behavior that will underlie their use of these tools. Models of behavior tend to be designed to suit particular environments being analyzed, and there is little debate over the proper core assumptions that should be held in common by such models. Agents are assigned decision rules and strategies without much of an attempt to link those rules to an explicit, much less unified, view of human nature. This is surprising in light of the fact that the hard sciences, much more than the social sciences, have traditionally been viewed as being governed by nomothetic approach to knowledge (Windelband, 1894), i.e. one that is aimed at creating general laws of nature that apply across environments and circumstances.

It is here that great potential exists for fruitful dialogue between hard and social scientists. The base assumptions of the rational choice model provide by far the most widely tested platform on which to build a general model of human behav-

ior, and can easily be adapted to many if not all of the agent-based modeling techniques and formalizations that have been developed over the last several years. However, the ongoing crisis of conventional rational choice provides an opportunity for hard scientists to not only to borrow the social scientist's preeminent general model, but also to help improve it by using their technical sophistication to suggest formal solutions to the issue of culture that will fit the criteria of generality and decisiveness, yet also be able to account for patterns found in behavioral data across environmental domains. In a way, this will return us full circle to the days of von Neumann and Morgenstern, when the boundaries between the hard and social science intellectual communities seemed more permeable than they have since.

References

1. Axelrod, Robert (2006 [1984]) *The Evolution of Cooperation*. Perseus Books, New York.
2. Axelrod, Robert (1997) *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton University Press, Princeton NJ.
3. Bainbridge, William Sims, Edward E. Brent, Kathleen M. Carley, David R. Heise, Michael W. Macy, Barry Markovsky, and John Skvoretz (1994) *Artificial Social Intelligence*. *Annual Review of Sociology* 20: 407-436.
4. Barry, Brian (1970) *Sociologists, Economists and Democracy*. Collier-MacMillan, London.
5. Camerer, Colin F., George Loewenstein, and Matthew Rabin, ed. (2003) *Advances in Behavioral Economics*. Princeton University Press, Princeton NJ.
6. Chai, Sun-Ki (1997) *Rational Choice and Culture: Clashing Perspectives or Complementary Modes of Analysis?* In: Richard Ellis and Michael Thompson, ed., *Culture Matters* 45-56. Westview Press, Boulder, CO.
7. Chai, Sun-Ki (2001) *Choosing an Identity: A General Model of Preference and Belief Formation*. University of Michigan Press, Ann Arbor, MI.
8. Chai, Sun-Ki and Brendon Swedlow, Ed. (1998) *Culture and Social Theory*. Transaction Publishers, New Brunswick NJ.
9. Chai, Sun-Ki and Aaron Wildavsky (1994) *Cultural Theory, Rationality and Political Violence*. In: Dennis J. Coyle and Richard J. Ellis, Eds. *Politics, Culture and Policy: Applications of Cultural Theory*. Westview Press, Boulder.
10. Douglas, Mary (1970) *Natural Symbols: Explorations in Cosmology*. Pantheon, New York.
11. Douglas, Mary (1982) *In the Active Voice*. Routledge and Keegan Paul, London.
12. Fudenberg, Drew, and Eric Maskin (1986) *The Folk Theorem in Repeated Games with Discounting or with Incomplete Information*. *Econometrica* 54:533-54.
13. Fudenberg, Drew and David K. Levine (1998) *The Theory of Learning in Games*. MIT Press, Cambridge, MA.
14. Gintis, Herbert, Samuel Bowles, Robert T. Boyd, and Ernst Fehr (2006) *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. MIT Press, Cambridge MA.
15. Harsanyi, John (1969) *Rational Choice Models of Political Behavior vs. Functionalist and Conformist Theories*. *World Politics* 21 (June):513-38.
16. Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis (2004) *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford University Press, New York.

17. Lazear, Edward P. (2000) Economic Imperialism. *Quarterly Journal of Economics* 115,1 (Feb):99-146.
18. Levitt, Steven D. and Stephen J. Dubner (2006) *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. William Morrow, New York.
19. Lepenies, Wolf and R. J. Hollingdale (1988) *Between Literature and Science: The Rise of Sociology*. Cambridge University Press, Cambridge.
20. Macy, Michael W. and Robert Willer (2002). From Factors To Actors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology* 28, 143-66.
21. Maynard Smith, John (1982) *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
22. Mirrowski, Philip (1991) *More Heat than Light: Economics as Social Physics, Physics as Nature's Economics*. Cambridge University Press, Cambridge.
23. Nelson, Richard R. and Sidney G. Winter (2006) *An Evolutionary Theory of Economic Change*. Belknap Press, Cambridge, MA.
24. Parthasarathy, T., B. Dutta, J.A.M. Potters, T.E.S. Raghaven, D. Ray, A. Sen (1997) *Game Theoretical Applications to Economics and Operations Research*. Springer, Berlin.
25. Radnitzky, Gerard, and Peter Bernholz, eds. (1987) *Economic Imperialism: The Economic Approach Applied Outside the Field of Economics*. Paragon House, New York.
26. Shubik, Martin (2002) Game Theory and Operations Research: Some Musings 50 Years Later. *Operations Research* 50,1 (Jan/Feb):192-196.
27. Thompson, Michael, Richard Ellis and Aaron Wildavsky (1990) *Cultural Theory*. Westview Press, Boulder CO.
28. von Neumann, John, and Oskar Morgenstern (1944) *Theory of Games and Economic Behavior*. Princeton University Press, Princeton NJ.
29. Weibull, Jörgen W. (1997) *Evolutionary Game Theory*. MIT Press, Cambridge, MA.
30. Windelband, Wilhelm (1894) *Geschichte und Naturwissenschaft*, Heitz, Strassburg